[19]中华人民共和国国家知识产权局

[51]Int.Cl[6]
G06F 12/00

[12]发明专利申请公开说明书

[21]申请号 98116228.2

[43]公开日 1999年2月24日

[11]公开号 CN 1208891A

[22]申请日 98.8.7 [21]申请号 98116228.2
[30]优先权
[32]97.8.8 [33]JP [31]214656/97
[71]申请人 株式会社东芝
地址 日本神奈川县
[72]发明人 关户一纪

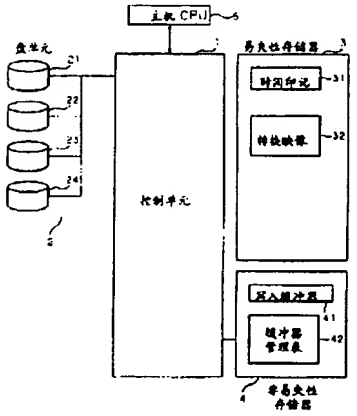
[74]专利代理机构 中国专利代理(香港)有限公司
代理人 杨 凯 叶恺东

权利要求书5页 说明书16页 附图页数14页

[54]发明名称 盘存储装置的数据更新方法和盘存储控制装置

[57]摘要

本发明提出一种原理上不需要间接变换映象、便宜且高速的盘存储装置的数据更新方法，并构成一种实现该方法的盘存储控制系统。在由N台盘装置构成的盘存储装置中，备有具有与N×K(整数)个逻辑块相当的容量的写入缓冲器，把应更新数据的逻辑块存储在该写入缓冲器中，控制装置1使该逻辑块的更新延迟到该已存储的逻辑块达到N×K-1个为止，把N×K-1个逻辑块并加上逻辑地址标记块的N×K个逻辑块依次连续写入N台盘装置上的空闲区域中。



权 利 要 求 书

1. 一种盘存储装置的数据更新方法，该盘存储装置具有N台盘装置（2）和根据主机（5）的命令向上述N台盘装置（2）写入数据或从上述N台盘装置（2）读出数据的控制装置（1），其特征在于：上述盘存储装置包括与上述控制装置（1）连接的易失性存储装置（3）和非易失性存储器（4），易失性存储装置（3）包含时间印记存储部（31）和变换映象存储部（32），非易失性存储器（4）包含具有与N×K（整数）个逻辑块的数据相当的存储容量的写入缓冲存储部（41）和缓冲管理表存储部（42），将应更新的逻辑块的数据存储到上述写入缓冲存储部（41）中，直到逻辑块个数达到N×K－1块为止，同时生成包含对于这些各逻辑块的逻辑地址和存储在时间印记存储部（31）中的时间印记的逻辑地址标记块，将它附加在上述N×K－1个逻辑块上，总共是N×K个逻辑块，依次将它们连续地写入上述N台盘装置（2）上的与分别保持上述应被更新的数据的逻辑地址区不同的别的空闲地址区中。

2. 如权利要求1所述的盘存储装置的数据更新方法，其特征在于：上述写入是在横跨多个盘存储装置的条形区中进行写入。

3. 如权利要求2所述的盘存储装置的数据更新方法，其特征在于：每当将上述写入缓冲器（41）中存储的N×K个逻辑块写入上述N台盘装置（2）时，使上述时间印记存储部（31）的值递增。

4. 如权利要求3所述的盘存储装置的数据更新方法，其特征在于：读出记录在上述盘装置的各条形区的逻辑地址标记块，通过检查该逻辑地址标记块检测出与各逻辑地址对应的盘装置上的物理存储位置，对检测出的存储位置进行写入或读出。

5. 如权利要求4所述的盘存储装置的数据更新方法，其特征在于：在上述逻辑地址标记块的检查中，当有多个包含相同逻辑地址的条形区时，将上述逻辑地址标记块内的时间印记是最新的一条逻辑地址块作为有效块，把其它条形区的具有相同逻辑地址的块判定为无效块。

6. 如权利要求4所述的盘存储装置的数据更新方法，其特征在于：在上述逻辑地址标记块的检查中，寻找最大时间印记值，再生由下一次写入所附加的时间印记。

7. 如权利要求4所述的盘存储装置的数据更新方法，其特征在于：在上述逻辑地址标记块的检查中，寻找最小时间印记值，求出作为

写入顺序判定基准的时间印记值.

8. 如权利要求 5 所述的盘存储装置的数据更新方法，其特征还在于：读出存储在上述盘存储装置中的多个条形区的逻辑块的数据，通过检查上述逻辑地址标记块，只将各条形区内的有效逻辑块送到上述写入缓冲器，生成与这些有效逻辑块对应的新的逻辑地址标记块，将由已送入上述写入缓冲存储部（41）的有效数据和新生成的逻辑地址标记构成的 1 条逻辑块依次写入与上述读出的多个条形区不同的另一个空闲区域中，由此，在上述盘存储装置上形成能够连续地写入逻辑块的空闲区域.

9. 如权利要求 8 所述的盘存储装置的数据更新方法，其特征在于：在生成上述新的逻辑地址标记块时，当有效块的个数不到 N×K－1 个时，对与不存储新的逻辑地址标记块内的数据的块对应的逻辑地址设定无效地址.

10. 如权利要求 8 所述的盘存储装置的数据更新方法，其特征在于：在把数据写入上述盘装置的空闲区域之后，在启动上述盘装置时检查上述多个条形区的逻辑地址标记块，据此，将与判断为有效的各逻辑地址对应的条号、条内的块号和有效数据的时间印记记录到上述变换映象中.

11. 如权利要求 10 所述的盘存储装置的数据更新方法，其特征在于：在作成上述变换映象的记录之后，在向盘装置进行的存取少的时间段内读出各条的逻辑地址标记块并与变换映象进行比较修正.

12. 如权利要求 4 所述的盘存储装置的数据更新方法，其特征在于：根据条对记录逻辑地址标记块的盘装置进行分散配置，在检查逻辑地址标记块时，将不同盘装置的逻辑地址标记块并列地读出.

13. 如权利要求 4 所述的盘存储装置的数据更新方法，其特征在于：上述逻辑地址标记与逻辑块数据一起依次写入各条形区，同时，还并列地写入专用标记区，在检查上述逻辑地址标记块时依次读该专用标记区并进行检查.

14. 如权利要求 4 所述的盘存储装置的数据更新方法，其特征在于：将盘装置上的存储区分割成以多个条为单位的多个段，控制成在一定时间内将条的数据只能写入 1 个段内，同时，在切换写入对象段时，将该时刻的上述变换映象（32）的内容和切换目标段的号码记录在盘装

置内，在作成以后的变换映象（32）时，只检查上述段切换时的变换映象的内容和写入已记录在上述盘装置中的段号码内的条形区域的逻辑地址标记块.

15. 如权利要求 4 所述的盘存储装置的数据更新方法，其特征在于：在上述非易失性存储器上准备与上述段内的各条对应的位映象，在切换写入对象段时，清除该位映象，在向条形区写入数据时，将与已写好的条形区对应的位进行置位，在作成变换映象（32）时，只检查在盘装置的段切换时的变换映象和已记录在上述盘装置（2）的段号的逻辑地址标记中已将上述位映象置位的逻辑地址标记.

16. 如权利要求 7 所述的盘存储装置的数据更新方法，其特征在于：为了寻找时间印记的最小值，即使对无效块少的条也定期进行读出，只将有效块送人上述写入缓冲器，从对应的逻辑标记块的逻辑地址和新的时间印记生成逻辑地址标记块，把由写入缓冲器的有效数据和已生成的逻辑地址标记块构成的条依次写入与保存有已读出的条的区域不同的另外的空闲区中.

17. 如权利要求 4 所述的盘存储装置的数据更新方法，其特征在于：为了寻找时间印记的最小值，对无效块少的条定期地只读出逻辑地址标记块，生成附加了已把无效块的逻辑地址作为 NULL 地址的新的时间印记的逻辑地址标记块，把在这里生成的逻辑地址标记块写在已读出的逻辑地址标记块的上面.

18. 如权利要求 10 所述的盘存储装置的数据更新方法，其特征在于：在变换映象（32）作成后，和与盘装置上的逻辑地址标记块的时间印记对应的变换映象（32）的时间印记进行比较，判定无效块.

19. 一种由 N 台盘装置（2）构成的盘存储装置的数据更新方法，其特征在于：上述盘存储装置（2）备有具有相当于（N—1）×K 个逻辑块的容量的写入缓冲器（41），把应更新数据的逻辑块存储在该写入缓冲器（41）中，使该逻辑块的更新延迟到该已存储的逻辑块达到所选择的个数，生成由对于上述写入缓冲器（41）存储的各逻辑块的逻辑地址构成的逻辑地址标记块，由在所选择的个数的逻辑块中附加了上述逻辑地址标记块的（N—1）×K 个数据逻辑块生成 K 个奇偶块，通过连续的写入工作将在该数据逻辑块中附加了奇偶块的 N×K 个逻辑块依次写入 N 台盘装置上（2）的与保存了应被更新的数据的区域不同的另外的空闲

区域中.

20. 如权利要求 19 所述的盘存储装置的数据更新方法，其特征在于：使上述选择的个数为（N－1）×K－1，以便在 1 个盘装置上记录逻辑地址标记块.

21. 如权利要求 19 所述的盘存储装置的数据更新方法，其特征在于：使上述选择的个数为（N－1）×K－2，分配 2 个逻辑地址标记块，以便用 1 个奇偶条在 2 个盘装置上记录逻辑地址标记块.

22. 如权利要求 20 所述的盘存储装置的数据更新方法，其特征在于：在检查记录在盘装置上的逻辑地址标记块方面，除了奇偶条单位的依次写入之外，还将该逻辑地址标记写入集中了逻辑地址标记的专用标记区，虽不用奇偶保护该专用标记区的写入数据，但使奇偶条内的记录逻辑地址标记的盘装置与记录专用标记区的逻辑地址标记的盘装置不同.

23. 一种盘存储控制装置，具有 N 台盘装置（2）和根据主机（5）的命令向上述 N 台盘装置（2）写入数据或从上述 N 台盘装置（2）读出数据的控制装置（1），其特征在于：具有与该盘存储控制装置的上述控制装置（1）连接的易失性存储装置（3）和与上述控制装置（1）连接的非易失性存储器（4），易失性存储装置（3）包含时间印记存储部（31）和变换映象存储部（32），非易失性存储器（4）包含具有与 N×K（整数）个逻辑块的数据相当的存储容量的写入缓冲存储部（41）和缓冲管理表存储部（42），将应更新的逻辑块的数据存储到上述写入缓冲存储部（41）中，直到逻辑块个数达到 N×K－1 块为止，同时生成包含对于这些各逻辑块的逻辑地址和存储在时间印记存储部（31）中的时间印记的逻辑地址标记块，将它附加在上述 N×K－1 个逻辑块上，总共是 N×K 个逻辑块，依次将它们连续地写入上述 N 台盘装置（2）上的与分别保持上述应被更新的数据的逻辑地址区不同的别的空闲地址区中.

24. 如权利要求 23 所述的盘存储控制装置，其特征在于：具有存储维持写入的时间顺序的时间印记的易失性存储器（3）、将应写入到盘装置上的数据变成记录表格的形式后保存的上述写入缓冲器（41）和存储写入缓冲器内的空闲区域及保存所保存的写入数据的逻辑地址信息的缓冲器管理信息的非易失性存储器（4）.

25. 一种盘存储控制装置，该装置具有由 N 台盘装置（2）构成的

盘存储装置，其特征在于：具有与（N－1）×K 个逻辑块相当的容量的写入缓冲器（41）和控制装置，所述控制装置把应更新数据的逻辑块存储在该写入缓冲器（41）中，使该逻辑块的更新延迟直到该已存储的逻辑块达到所选择的个数为止，生成由对于上述写入缓冲器（41）中已存储的各逻辑块的逻辑地址构成的逻辑地址标记块，从在选择个数的逻辑块中附加了上述逻辑地址标记块的（N－1）×K 个数据逻辑块生成 K 个奇偶块，通过连续的写入工作将在该数据逻辑块附加了奇偶的 N×K 个逻辑块依次写入 N 台盘装置上的与保存了应被更新的数据的区域不同的另外的空闲区域中.

26. 如权利要求 25 所述的盘存储控制装置，其特征在于：为了采用使用了奇偶检验的冗余性的盘结构而附加冗余盘装置，进而还具有存储维持写入的时间顺序的时间印记的易失性存储器（3）、将应写入到盘装置上的数据变成记录表格的形式后保存的上述写入缓冲器（41）和存储写入缓冲器内的空闲区域及保存所保存的写入数据的逻辑地址信息的缓冲器管理信息的非易失性存储器（4）.

说　明　书

盘存储装置的数据更新方法和盘存储控制装置

本发明涉及使用了 RAID（Redundant Array of Inexpensive Disk 廉价盘的冗余阵列）等多台盘存储装置的盘阵列存储装置，特别是涉及对盘阵列存储装置进行数据存储和读出的方法及其所用的控制装置．在美国专利第 5,124,987 号、特开平 6—214720 和特开平 6—266510 号所公开的 RAID 方式的存储装置中，作为向存储装置高速写入的方法，提出了一种改写数据的方法，即，不是改写旧数据区，而是将更新的数据整理后写入盘装置内的预先准备好的另外的空区域中，然后，在其后的空闲时间内再改写旧数据区．

使用图 18 简单地说明上述现有的方法．在图中，考虑更新已存储在逻辑块地址（以下仅称为逻辑地址）L6、L4、L2、L12、L7、L11 内的数据块的情况．这些逻辑块地址 L6、L4、L2、L12、L7、L11 内的旧数据存在于 3 个盘装置 181、182、183 内的物理块地址（以下仅称为物理地址）P6、P4、P2、P12、P7、P11 中．首先，应更新的新数据块 L6 数据、L4 数据、L2 数据、L12 数据、L7 数据、L11 数据通常暂时存储在由非易失性存储器构成的写入缓冲存储器 184 中．这些数据块不是直接去替换存储了要更新的旧数据的物理块地址 P6、P4、P2、P12、P7、P11 的内容、即数据，而是保持旧数据不变，将该更新的数据块整理后写入盘装置 181～183 内的预先准备好的另外的空区域、即物理地址 P51、P52、P53、P54、P55、P56 中．该写入工作是向 3 个盘装置 181、182、183 内的连续物理地址 P51—P52、P53—P54、P55—P56 写入的，所以，与直接进行替换时需要 6 次写入工作相比，减少到实际上只要 3 次写入工作，写入性能大大提高．

另一方面，在这种现有的盘阵列存储装置中，设有表示数据块与存储的逻辑地址和物理地址的对应关系的表、即间接映象（map）．在数据更新时，如上所述，逻辑地址 L6、L4、L2、L12、L7、L11 内的最新数据实际上是存在于盘装置内的物理地址 P51、P52、P53、P54、P55、P56 中，所以，改写间接映象的内容使它正确地指向盘上的位置．即，例如，逻辑地址 L6 内的数据块本来必须在盘装置 181 内的物理地址 P6 中，但实际上存储在物理地址 P51 内，所以，将与间接映象 175 内的逻辑地址

L6 对应的物理地址 P6 改写成 P51. 以下，同样分别将与间接映象 185 内的逻辑地址 L4、L2、L12、L7、L11 对应的物理地址改写成 P52、P53、P54、P55、P56.

此外，因为在将存储在盘阵列存储装置中的数据读出时，是求出与间接映象 185 所指定的逻辑地址对应的存储了最新数据块的物理地址后再读出的，所以没有将旧数据读出的危险。

再有，在图 18 所示的例子中，为使说明简单起见，作为存储的数据块，对 1 台盘装置只写入 2 个块的数据，但实际上要写入几十个数据块。

在上述现有的技术中，因为是通过间接映象去管理最新数据的位置信息，所以，存在当间接映象因故障或误工作而使其数据丢失时盘装置内的全部数据便丢失的所谓数据安全性问题. 此外，因为必须对全部逻辑块准备间接映象，而且当发生电源故障时还要保持间接映象，所以必需要大容量的非易失性存储器，因而存在间接映象非常贵的问题.

本发明是为了解决上述问题而提出的，其目的在于提供一种原理上不需要间接映象、便宜且快速的盘存储装置的数据更新方法以及盘存储控制系统。

本发明的盘存储装置的数据更新方法的特征在于，具有由 N 台盘装置、根据主机的命令向上述 N 台盘装置写入数据或从上述 N 台盘装置读出数据的控制装置、易失性存储器和非易失性存储器构成的盘存储装置，该易失性存储器与该控制装置连接，并包含时间印记（stamp）存储部和变换映象存储部，该非易失性存储器与该控制装置连接，并包含具有与 N×K（整数）个逻辑块的数据相当的存储容量的写入缓冲存储部和缓冲管理表存储部，将应更新的逻辑块的数据存储到上述写入缓冲器中，直到逻辑块个数达到 N×K—1 块为止，同时生成包含这些各逻辑块的逻辑地址和存储在时间印记存储部中的时间印记的逻辑地址标记（tag）块，将它附加在上述 N×K—1 个逻辑块上，总共是 N×K 个逻辑块，依次将它们连续地写入上述 N 台盘装置上的与分别保持上述应被更新的数据的逻辑地址区不同的别的空闲地址区中.

本发明的盘存储装置的数据更新方法的特征还在于，上述写入是在横跨多个盘存储装置的条形区中进行写入.

本发明的盘存储装置的数据更新方法的特征还在于，每当将上述写

入缓冲器中存储的 N×K 个逻辑块写入上述 N 台盘装置时，使上述时间印记存储部递增。

本发明的盘存储装置的数据更新方法的特征还在于，读出记录在上述盘装置的各条形区的逻辑地址标记块，通过检查该逻辑地址标记块检测出与各逻辑地址对应的盘装置上的物理存储位置，对检测出的存储位置进行写入或读出。

本发明的盘存储装置的数据更新方法的特征还在于，在上述逻辑地址标记块的检查中，当有多个包含相同逻辑地址的条形区时，将上述逻辑地址标记块内的时间印记是最新的一条逻辑地址块作为有效块，把其它条形区的具有相同逻辑地址的块判定为无效块。

本发明的盘存储装置的数据更新方法的特征还在于，在上述逻辑地址标记块的检查中，寻找最大时间印记值，再生由下一次写入所附加的时间印记。

本发明的盘存储装置的数据更新方法的特征还在于，在上述逻辑地址标记块的检查中，寻找最小时间印记值，求出作为写入顺序判定基准的时间印记值。

本发明的盘存储装置的数据更新方法的特征还在于，读出存储在上述盘存储装置中的多个条形区的逻辑块的数据，通过检查上述逻辑地址标记块，只将各条形区内的有效逻辑块送到上述写入缓冲器，生成与这些有效逻辑块对应的新的逻辑地址标记块，将由已送入上述写入缓冲器的有效数据和新生成的逻辑地址标记构成的 1 条逻辑块依次写入与上述读出的多个条形区不同的另一个空闲区域中，由此，在上述盘存储装置上形成能够连续地写入逻辑块的空闲区域。

本发明的盘存储装置的数据更新方法的特征还在于，在生成上述新的逻辑地址标记块时，当有效块的个数不到 N×K－1 个时，对与不存储新的逻辑地址标记块内的数据的块对应的逻辑地址设定无效地址。

本发明的盘存储装置的数据更新方法的特征还在于，在把数据写入上述盘装置的空闲区域之后，在启动上述盘装置时检查上述多个条形区的逻辑地址标记块，据此，将与判断为有效的各逻辑地址对应的条号、条内的块号和有效数据的时间印记记录到上述变换映象中。

本发明的盘存储装置的数据更新方法的特征还在于，在作成上述变换映象的记录之后，在向盘装置进行的存取少的时间段内读出各条的逻

辑地址标记块并与变换映象进行比较修正。

本发明的盘存储装置的数据更新方法的特征还在于，根据条对记录逻辑地址标记块的盘装置进行分散配置，在检查逻辑地址标记块时，将不同盘装置的逻辑地址标记块并列地读出。

本发明的盘存储装置的数据更新方法的特征还在于，上述逻辑地址标记与逻辑块数据一起依次写入各条形区，同时，还并列地写入专用标记区，在检查上述逻辑地址标记块时依次读该专用标记区并进行检查。

本发明的盘存储装置的数据更新方法的特征还在于，将盘装置上的存储区分割成以多个条为单位的多个段，控制成在一定时间内将条的数据只能写入 1 个段内，同时，在切换写入对象段时，将该时刻的上述变换映象的内容和切换目标段的号码记录在盘装置内，在作成以后的变换映象时，只检查上述段切换时的变换映象的内容和写入已记录在上述盘装置中的段号码内的条形区域的逻辑地址标记块。

本发明的盘存储装置的数据更新方法的特征还在于，在上述非易失性存储器上准备与上述段内的各条对应的位映象，在切换写入对象段时，清除该位映象，在向条形区写入数据时，将与已写好的条形区对应的位进行置位，在作成变换映象时，只检查在盘装置的段切换时的变换映象和已记录在上述盘装置的段号的逻辑地址标记中已将上述位映象置位的标记。

本发明的盘存储装置的数据更新方法的特征还在于，为了寻找时间印记的最小值，即使对无效块少的条也定期进行读出，只将有效块送人上述写入缓冲器，从对应的逻辑标记块的逻辑地址和新的时间印记生成逻辑地址标记块，把由写入缓冲器的有效数据和已生成的逻辑地址标记块构成的条依次写入与保存有已读出的条的区域不同的另外的空闲区中。

本发明的盘存储装置的数据更新方法的特征还在于，为了寻找时间印记的最小值，对无效块少的条定期地只读出逻辑地址标记块，生成附加了已把无效块的逻辑地址作为 NULL 地址的新的时间印记的逻辑地址标记块，把在这里生成的逻辑地址标记块写在已读出的逻辑地址标记块的上面。

本发明的盘存储装置的数据更新方法的特征还在于，在变换映象作成后，和与盘装置上的逻辑地址标记块的时间印记对应的变换映象的时

间印记进行比较，判定无效块。

本发明的盘存储装置的数据更新方法的特征还在于，在由 N 台盘装置构成的盘存储装置中，备有具有相当于（N－1）×K 个逻辑块的容量的写入缓冲器，把应更新数据的逻辑块存储在该写入缓冲器中，使该逻辑块的更新延迟到该已存储的逻辑块达到所选择的个数，生成由上述写入缓冲器存储的各逻辑块的逻辑地址构成的逻辑地址标记块，由在所选择的个数的逻辑块中附加了上述逻辑地址标记块的（N－1）×K 个数据逻辑块生成 K 个奇偶块，通过连续的写入工作将在该数据逻辑块中附加了奇偶块的 N×K 个逻辑块依次写入 N 台盘装置上的与保存了应被更新的数据的区域不同的另外的空闲区域中。

本发明的盘存储装置的数据更新方法的特征还在于，使上述选择的个数为（N－1）×K－1，以便在 1 个盘装置上记录逻辑地址标记块。

本发明的盘存储装置的数据更新方法的特征还在于，使上述选择的个数为（N－1）×K－2，分配 2 个逻辑地址标记块，以便用 1 个奇偶条在 2 个盘装置上记录逻辑地址标记块。

本发明的盘存储装置的数据更新方法的特征还在于，在检查记录在盘装置上的逻辑地址标记块方面，除了奇偶条单位的依次写入之外，还将该逻辑地址标记写入集中了逻辑地址标记的专用标记区，虽不用奇偶保护该专用标记区的写入数据，但使奇偶条内的记录逻辑地址标记的盘装置与记录专用标记区的逻辑地址标记的盘装置不同。

本发明的盘存储控制装置的特征在于，具有由 N 台盘装置、根据主机的命令向上述 N 台盘装置写入数据或从上述 N 台盘装置读出数据的控制装置、易失性存储器和非易失性存储器构成的盘存储装置，该易失性存储器与该控制装置连接，并包含时间印记存储部和变换映象存储部，该非易失性存储器与该控制装置连接，并包含具有与 N×K（整数）个逻辑块的数据相当的存储容量的写入缓冲存储部和缓冲管理表存储部，将应更新的逻辑块的数据存储到上述写入缓冲器中，直到逻辑块个数达到 N×K－1 块为止，同时生成包含这些各逻辑块的逻辑地址和存储在时间印记存储部中的时间印记的逻辑地址标记（tag）块，将它附加在上述 N×K－1 个逻辑块上，总共是 N×K 个逻辑块，依次将它们连续地写入上述 N 台盘装置上的与分别保持上述应被更新的数据的逻辑地址区不同的别的空闲地址区中。

本发明的盘存储控制装置的特征在于，还具有存储维持写入的时间顺序的时间印记的易失性存储器、将应写入到盘装置上的数据变成记录表格的形式后保存的上述写入缓冲器和存储写入缓冲器内的空闲区域及保存所保存的写入数据的逻辑地址信息的缓冲器管理信息的非易失性存储器。

本发明的盘存储控制装置的特征在于，还具备由 N 台盘装置构成的盘存储装置、具有与（N—1）×K 个逻辑块相当的容量的写入缓冲器和控制装置，所述控制装置把应更新数据的逻辑块存储在该写入缓冲器中，使该逻辑块的更新延迟到该已存储的逻辑块达到所选择的个数为止，生成由对于上述写入缓冲器中已存储的各逻辑块的逻辑地址构成的逻辑地址标记块，从在选择个数的逻辑块中附加了上述逻辑地址标记块的（N—1）×K 个数据逻辑块生成 K 个奇偶块，通过连续的写入工作将在该数据逻辑块中附加了奇偶块的 N×K 个逻辑块依次写入 N 台盘装置上的与保存了应被更新的数据的区域不同的另外的空闲区域中。

本发明的盘存储控制装置的特征还在于，为了采用使用了奇偶检验的冗余性的盘结构而附加冗余盘装置，进而还具有存储维持写入的时间顺序的时间印记的易失性存储器、将应写入到盘装置上的数据变成记录表格的形式后保存的上述写入缓冲器和存储写入缓冲器内的空闲区域及保存所保存的写入数据的逻辑地址信息的缓冲器管理信息的非易失性存储器。

通过采用上述结构，原则上不需要间接映象，可以构筑一种便宜且高速的盘存储装置以及盘存储控制系统。

图 1 是表示本发明的实施例的框图。

图 2 是为显示本发明的实施例中写入缓冲器和缓冲器管理信息的关系所引用的图。

图 3 是表示本发明的实施例中盘装置的空闲区存储的内容的图。

图 4 是为显示从主机写入 1 个块的数据的写入顺序所引用的图。

图 5 是表示图 4 的例子中条 ST1、ST2 的逻辑地址标记 TG1/TG2 的内容的图。

图 6 是表示将条 ST3/ST4 统一成 1 个条 ST5 的例子的图。

图 7 是表示在进行条的统一时从逻辑地址标记 TG3/TG4 作成逻辑地址标记 TG5 的例子的图。

图 8 是表示本发明的实施例中使用的变换映象的结构例的图.

图 9 是用于说明系统启动时变换映象的作成方法的流程图.

图 10 是表示将根据条分散配置了 4 台存储逻辑地址标记的盘装置的例子的图.

图 11 是表示在段分割中的盘装置的存储区域的分配的图.

图 12 是表示段管理信息的输入项目结构的图.

图 13 是表示连续地存储逻辑地址标记的专用标记区的内容的图.

图 14 是表示应用本发明而构成的基于 RAID5 的盘装置的实施例的框图.

图 15 是表示图 13 所示的实施例的工作原理的图.

图 16 是表示控制成将相同的逻辑地址标记写入 2 个盘装置的例子的图.

图 17 是表示为了高速地作成变换映象而分配使用专用标记区的情况的例子的图.

图 18 是表示现有例中实现数据更新方法的系统结构的图.

图 1 使用本发明构成的盘存储装置的原理结构图. 本发明的盘存储装控制系统由控制装置 1、盘装置 2、易失性存储器 3 和非易失性存储器 4 构成. 盘装置 2 由多台盘存储装置构成, 但在该实施例中为说明简单起见, 以由 4 台 21、22、23、24 构成的盘装置为例进行说明. 易失性存储器 3 中设有存储写入的时间顺序的时间印记区 31 和间接映象存储区 32. 非易失性存储器 4 中设有将写入盘装置 2 的数据作成记录表的结构并保存的写入缓冲区 41 和保存写入缓冲区 41 中的空闲区域及已保存的写入数据的逻辑地址的信息的缓冲器管理表 42. 控制装置 1 根据主机 5 来的命令管理这些时间印记区 31、写入缓冲区 41 和缓冲器管理表 42, 控制向盘装置 2 的写入.

图 2 示出分配给非易失性存储器 4 的写入缓冲区 41 和缓冲器管理表 42 的关系. 控制装置 1 不是直接将与外部连接的主机所要求的写入数据写入盘装置 2, 而是将它们以块为单位分割后再顺序 (以记录表的形式) 存储到写入缓冲区 41 中. 在此, 缓冲器管理表 41 形成由多个输入项目构成的表格, 在这些各输入项目中, 与缓冲区 41 内的各块位置 B0、B1、…B15 对应地保存应写入的各数据块的从主机看到的逻辑地址. 在缓冲器管理表 42 内的各输入项目中还设立标志“F”, 表示该输入项

目已分配了数据，对没有分配数据的输入项目设立标志“U”.

在图 2 所示的例子中示出，写入数据存储在写入缓冲区 41 内的到块位置 B7 为止的区域中，B0、B1、…B7 的逻辑地址是 LA134、LA199、…LA678.

此外，以称之为条单元的存储区为单位（其长度接近于该盘装置的 1 个信道（track）的长度即可），对于盘装置 2 的各盘存储装置 21～24 进行数据的写入，该数据的长度相当于块长度的整数（K）倍、即 K 个块的长度. 而且，与各盘存储装置 21～24 的物理位置对应的条单元整体上作为 1 个条形区（ST），也以相同的时序进行写入.

此外，盘装置 2 向主机 2 呈现的存储容量比构成它的多台盘装置 21～24 加在一起的实际存储容量要小，即，当主机 5 最初询问存储容量时，作为回答返回的是较小的容量. 因此，除了从主机 5 可以进行逻辑读写的存储区域之外，还可确保多余的存储区域，即空闲区域.

进而，时间印记 31 是当从主机 5 来的写入数据实际上已写入盘装置 2 时所附加的信息，是为了判定向盘装置 2 写入数据的顺序而使用的. 因此，每当写入缓冲器 41 的数据写入盘装置 2 时，时间印记 31 就递增.

下面，参照图 2～图 8 详细说明图 1 所示的本发明实施例的工作.

首先，说明写入工作. 控制装置 1 从主机 5 接受应写入的数据及其逻辑地址，如图 2 所示，将数据以块为单位分割，按顺序连续地存储到非易失性存储器 4 的写入缓冲区 41 的空闲区域中. 再有，在图 2 中，依次连续地写入到写入缓冲区 41 的相当于由 B0、B1、…B15 形成的 15 个块长的空闲区内.

此外，将接受的逻辑地址变换成每个块的地址，并存储到与缓冲器管理表 42 对应的输入项目 B0、B1、…B15 中. 再有，当对已存储到写入缓冲器 41 的数据进行更新时，不是依次存储到写入缓冲器 41 的空闲区，而是直接改变写入缓冲器 41 的旧数据.

在来自主机 5 的写入数据存储在相对于盘装置 2 的 1 个条（ST）的区域只少 1 个块的数目的写入缓冲器 41 中、即存储在（4K－1）块的写入缓冲器 41 中时，控制装置 1 将这些数据写入盘装置 2. 在图 2 中，是在存储于于 K=4、即 15 个块的写入缓冲器 41 中时进行对于盘装置 2 的写入. 这时，作为最后的写入块，根据存储在缓冲器管理表 42 内的写入管理表中的各个块的逻辑地址和易失性存储器 3 上的时间印记 31 作

成图 3 所示那样的逻辑地址标记块 LA—TAG. 事先在该逻辑地址标记块的地址数据和数据块之间设一一对应的关系，就可以知道各数据块的逻辑地址.

然后，如图 3 所示，将附加了该逻辑地址标记块的 1 个条的数据整理后同时写入盘装置 21～24 的空闲区. 在图 3 中，盘装置 21～24 的 1 个条（ST）的空闲区由 4 个单元条 D1～D4 表示，图 3 示出了写入各单元条 D1～D4 区域内的 4 个数据块的逻辑地址. 再有，图 1 的时间印记 31 的值在写入结束时加 1. 这样，将很多零碎的盘写入工作归纳成 1 次写入，所以大大地提高了盘的写入性能.

其次，说明数据块的重新装入处理. 作为本发明的盘写入方法，不是直接改写旧的数据区，而是将更新的数据存起来，归纳后写入盘装置 2 的事先准备好的另外的空闲区域中. 在这样的盘写入方法中，必须始终存在用于将存储在盘装置 2 内的数据归纳后再写入的空闲区. 因此，可以在没有进行从主机 5 来的盘的存取的空闲时间内把已写入其它区域而变成无效的数据集中起来，作成空闲区. 将该处理称为重新装入处理. 该重新装入处理由判定无效块和条统一两个步骤构成.

作为判定无效块的例子，考虑有按图 4 所示的顺序从主机 5 写入 1 个块长的数据的情况. 图中 L××表示从主机送过来的逻辑地址，S××表示写入顺序. 在本发明的实施例中，因写入缓冲器 41 保存有 15 个块的数据，将最初 S1～S15 写入的数据归纳成 1 个条（ST1），附加时间印记 TS1 后写入盘装置的空闲区. 同样，S16～S30 的写入数据作为另外的条（ST2）并附加时间印记 TS2 后写入另外的空闲区. 再有，因每写入 1 次时间印记 31 加 1，故存在 TS1<TS2 的关系.

这里，由图可知，逻辑地址 L9、L18 的数据在时间印记为 TS1 的条中作为 S5、S2 的块、在时间印记为 TS2 的条中作为 S19、S21 的块重复存在. 即，存在 2 个应写入同一逻辑地址 L9、L18 的数据. 但是，若考虑写入数据块的顺序，后来写入的 S19、S21 的数据块是有效的，因此，S5、S2 的数据必须判定为无效.

然而，在此为方便而使用的写入顺序 S××在实际的盘上没有记录下来. 因此，使用附加在各个条中的逻辑地址标记进行判定. 在图 4 的例子中，2 个条 ST1、ST2 的逻辑地址标记 TG1、TG2 的内容如图 5 所示. 即，逻辑地址 TG1、TG2 将各块的逻辑地址存储在与写入到写入缓冲器 41

的 15 个块 B0、B1、…B15 对应的存储区中，在第 16 号存储区内分别写入条 ST1、ST2 的每一个被写入时的时间印记 TS1、TS2.

由图可知，2 个逻辑地址标记 TG1、TG2 包含相同的逻辑地址 L9、L18 的数据，条 ST1 的块 B5、B2 和条 ST2 的块 B4、B6 中的某一个的数据为无效数据. 进而，若将逻辑地址标记 TG1 的时间印记 TS1 与逻辑地址标记 TG2 的时间印记 TS2 进行比较，根据 TS1<TS2 的关系可以判定条 ST1 的块 B5、B2 为无效. 如上面说明的那样，通过调查盘装置 2 的逻辑地址就可以找出无效的数据块.

图 6 是表示条统一的例子的图，示出了将 2 个条 ST3、ST4 统一成 1 个条 ST5 的情况. 在该图中，对于条 ST3，假定 B2、B7、B8、B12、B13 的 5 个块有效，其余 10 个块无效（画阴影线）. 同样，对于条 ST4，假定 B18、B19、B20、B21、B22、B24、B25、B27、B29 的 9 个块有效，其余 6 个块无效（画阴影线）. 因此，在 2 个条 ST3、ST4 中，有效块加起来只有 14 块，通过只将这 2 个条 ST3、ST4 的有效块取出来统一变成 1 个条 ST5，结果作成相当于 1 个条的空闲区.

条统一的具体方法是，从易失性存储器 3 读出图 6 所示的 2 个条 ST3、ST4，只将这 2 个条 ST3、ST4 的有效块取出并将它们连续地转送到写入缓冲器 41 中. 与此相应，逻辑地址标记也如图 7 所示从 TG3、TG4 转移到只与有效块的逻辑地址对应的位置上，作成新的逻辑地址标记 TG5，将这时的时间印记更新为 ST5.

在该例中，因只有 14 个有效块，故进而等待从主机 5 供给 1 个写入块，使之凑成 1 个条，归纳后写入盘装置 2 的空闲区. 这时，虽然有效地利用了盘区域，但因为要等待从主机 5 供给写入块，故存在使向盘进行的存取集中在一起的危险. 因此，也可在存取的空隙时间按原有的空状态写入最后的数据块. 这时，由于通过在与逻辑地址标记 TG5 的最后数据块对应的逻辑地址上输入－1 等 NULL 地址来表示数据没有输入，所以不成为问题.

其次，说明这样写入的数据块的读出工作. 通过对盘装置 2 的全部条的逻辑地址标记进行重新装入处理的无效块判定，可以检测出对全部逻辑地址有效的块的物理位置. 因此，从原理上讲，通过在从主机 5 接受读出块的逻辑地址时进行全部条的检查，可以找出应读出的物理块. 但是，该方法在块读出时要耗费很长的时间，所以不实用.

因此，只在系统启动时进行全部条的逻辑地址标记的调查，在图 1 所示的易失性存储器 3 上根据逻辑地址作成向物理地址变换的映象 32. 使用该变换映象 32 对主机 5 来的读出请求进行向有效块的存取. 因此，可不对每一个主机 5 来的读出请求都进行地址标记的调查，不会大大降低读出时的性能. 此外，该变换映象 32 无论何时都能通过调查全部条来再生，所以，不需要象以往的间接映象那样为了防备电源故障而必须将其存储在非易失性存储器中.

在此，使用图 8 说明变换映象 32. 如图所示，变换映象 32 以表格的形式保存存储了相对各逻辑地址 L0 ~ Ln 的块的条号 ST#、该条内的块号 BLK3 及其时间印记 TS#. 因此，若给出逻辑地址 L0 ~ Ln，通过检索该表就可以简单地从 ST#和 BLK#求出实际的物理地址.

此外，关于系统启动时变换映象的作成，如图 9 的流程图所示那样，依次读出存储在盘装置 2 中的条 ST1、ST2、ST3、…的逻辑地址标记 TG1、TG2、TG3、…（步骤 1）. 将这样读出的逻辑地址标记 TG1、TG2、TG3、…内的各时间印记 TS1、TS2、TS3、…取出（步骤 2）. 进而，将读出的各逻辑地址标记 TG1、TG2、TG3、…内的各逻辑地址 LA × × 依次取出（步骤 3）. 当将与取出的逻辑地址 LAn 相同的逻辑地址 LAn 登录在变换映象 32 内时，将登录在变换映象 32 内的逻辑地址 LAn 的时间印记 TSn 与取出的逻辑地址 LAn 的时间印记 TSm 进行比较（步骤 4）. 作为该比较的结果，当取出的逻辑地址 LAn 的时间印记 TSi 比登录在变换映象 32 内的逻辑地址 LAn 的时间印记 TSj 新、即 TSi 大于 TSj 时，将存储了已作为登录在对于逻辑地址 LAn 的变换映象 32 内的条号 ST#取出的逻辑地址 LAn 的盘装置 2 的条号 ST # 存储起来，此外，还将存储了该逻辑地址 LAn 的条 ST#内的数据块的位置 BLK#存储起来（步骤 5）. 进而，将已作为对于变换映象 32 内的逻辑地址 LAn 的时间印记 TS#取出的、对于逻辑地址 LAn 的时间印记 TS#存储起来（步骤 6）. 作为上述步骤 4 的比较结果，当取出的逻辑地址 LAn 的时间印记 TSi 比登录在变换映象 32 内的逻辑地址 LAn 的时间印记 TSj 旧、即 TSi 小于 TSj 时，使对于登录在变换映象 32 内的逻辑地址 LAn 的条 ST#、数据块的位置 BLK#和时间印记 TS#的内容维持不变，针对已读出的逻辑地址标记内的所有的逻辑地址，调查是否执行了同样的处理（步骤 7）. 在对全部逻辑地址未执行同样的处理的情况下，返回步骤 3，重复执行直到步骤 8 的处理. 在对已读出的逻辑地

址标记内的所有的逻辑地址已完成同样的处理的情况下，则接着针对存储在盘装置 2 上的所有的逻辑地址标记 TG1、TG2、TG3、…，调查是否已执行上述处理（步骤 8）. 若未完成同样的处理，则返回步骤 3，重复执行直到步骤 7 的处理. 若已完成同样的处理，则使对于该时刻残留的逻辑地址的条 ST#、数据块的位置 BLK#和时间印记 TS#的内容变成变换映象 32 的登录内容（步骤 9）.

即，对于已取出的逻辑地址标记内的全部逻辑地址，只有当逻辑地址标记的时间印记比变换映象 32 内表格的时间印记大时，才将与该条号对应的块号登录在表格中. 若对全部条进行该调查，则可以作成只指示有效块的变换映象. 进而，每当向盘装置 2 写入条时，通过对该逻辑地址标记也进行同样的处理，只将始终有效的块登录在该变换映象 32 上. 此外，通过在盘存取的空隙时间内将各条的逻辑地址标记与变换映象进行比较检查，即使因存储器故障使该变换映象出现不正确的值，也能够检测出来并加以改正.

如上所述，作成变换映象的主要的处理是逻辑地址标记的检查. 所以，在象大容量盘装置那样逻辑地址标记数多时，作成电源故障和系统启动时的变换映象需要很长时间. 特别是，如图 2 所示，当逻辑地址标记块集中在 1 台盘装置 24 上时，在系统启动时存取工作集中在该盘上进行，不能平行地进行逻辑地址标记的调查. 因此，如图 10 所示，通过将根据条来存储逻辑地址标记的盘装置分散成 4 台平行地进行逻辑地址标记的调查，可以使作成该映象的时间缩短到 1/4.

此外，通过将盘装置 2 的存储区域分割成多个段进行管理，可以削减作成变换映象所必需的逻辑地址标记的检查个数. 图 11 示出段分割方式下盘装置的存储区域的结构. 如图所示，盘装置的存储区域以条为单位被分割成段管理信息部分（阴影部分）和 4 个段. 在此，所谓段是指写入缓冲器数据的一并写入和重新装入处理的盘写入集中在某一时间进行的单位区域. 例如，控制空闲区的选择，使得在段 2 是盘写入的对象的期间内不向段 1、3、4 进行写入.

此外，当某段的空闲区少、将盘写入切换到其它段时，将段管理信息保存在盘装置上. 段管理信息如图 12 所示那样由段号和切换时的变换映象构成. 所谓段号是指切换目标段的段号码，所谓切换时的变换映象是指段切换时刻的易失性存储器 3 上的变换映象的状态.

再有，每当段切换时，切换时的变换映象不是全都写在上面，只要返回写到当前段中已写入的逻辑地址的输入项目中即可．因此，通过在上一次段切换时记住时间印记并与变换映象的时间印记比较，就可以判定写入到当前段中的逻辑地址．

在该段分割方式中，在段切换时保存了段管理信息．因此，从段管理信息读出段切换时的变换映象，然后，只要检查由段管理信息的段号所指定的段的逻辑地址标记，就能够再现与检查全部逻辑地址标记的情况相同的变换映象．所以，利用该方式所必需的逻辑地址标记的检查数只要检查 1 个段的即可，在本例中作成变换映象所要的时间缩短到 1/4．

进而，在非易失性存储器 4 上准备好与段内全部条对应的位映象，在段切换时清除该位映象，在一并写入和重新装入时将与已写入的条对应的比特置位成“1”．由此，在段切换之后只是有变化的条的位映象变成“1”．因此，在作成变换映象时，通过参照该位映象、只检查有变化的条的逻辑地址标记，可使检查数进一步减少，使作成变换映象所要的时间进一步缩短．

通常逻辑地址标记的长度是 512～1024 字节．盘的顺序存取和随机存取大约有 50 倍的性能差．在图 2 所示的方式中，逻辑地址标记的信息对各个条是分散存在的，所以，进行的是在变换映象作成时很耗费时间的随机存取．因此，如图 13 所示那样，准备了专用标记区（在段分割的情况下，是对每个段准备的），该标记区只连续地存储逻辑地址标记，能以高到 50 倍速度的顺序存取来读出逻辑地址标记．

而且，在将主机来的数据一并写入或重新装入数据的写入时，不仅将逻辑地址标记写入空闲区还写入对应的专用标记区．在该方法中，在图 2 的方式下，每一个条有 4 次盘写入，由于向专用区写入逻辑地址而增加 1 次．但是，作成变换映象的速度却提高了 50 倍，所以，当在盘装置的建立时间方面出问题时，它是非常有效的方法．为了使向专用标记区的写入时间最少，使专用标记区如图 13 所示那样成为对象区域的中心，减少搜索时间．此外，盘装置 2 是以扇区（512 字节等）为单位写入的，专用标记区内的逻辑地址标记没有必要以扇区为单位进行分配并在逻辑地址标记写入时读出．

最后，就时间印记进行说明．如图 1 所示，因时间印记存储在易失性存储器 3 上，故由于电源故障等原因，易失性存储器 3 上的时间印记

会丢失．因此，与变换映象一样，只在系统启动时调查全部条的逻辑地址标记，使最大的时间印记 31 的下一个值置成易失性存储器 3 上的时间印记 31．再有，在作成变换映象的说明中所述的缩短时间的方法照样可以适用于时间印记的再生．

此外，每当写入盘装置时，时间印记 31 加 1，只在判定盘上的写入顺序时才使用．作为例子，说明时间印记 31 由 24 位计数器构成时的情况．在 24 位计数器中，计数器在 16M 次写入后完成 1 个循环而回到 0．因此，一般来说，以有效时间印记的最小值为基准，将比它小的值加 16M 后进行比较判定．该最小值也是一样，只在系统启动时才调查全部条的逻辑地址标记并将它求出来．

但是，可以使用该方法的前提是，时间印记的最大值没有超过最小值，即时间印记最大值和最小值的差是在能用 24 位表示的范围内．因此，时间印记 31 必需在 1 个循环前更新全部条并重新更新时间印记值．因此，即使无效块少也控制成将在预先设定的写入次数间没有被更新的条作为重新装入的对象选出，或者只改写将无效块的逻辑地址作为 NULL（无效）地址的条的逻辑地址标记．使用 NULL 地址的方法因为是改写逻辑地址标记块，故与重新装入相比是非常简单的处理。

再有，在上述实施例中，对于无效块的判定，只说明了将 2 个条 ST1、ST2 的逻辑地址标记相互比较来进行判定的方法，但要调查全部无效块则必需调查 2 个条之间的全部组合．然而，如果有变换映象，则可以对逻辑地址标记内的各逻辑地址将指示有效数据的变换映象的时间印记与该条的时间印记进行比较，将时间印记值小的块判定为无效块．

图 1 已示出将数据分散在多个盘上的 RAID0 的结构，但本发明的方式也可以适用于使用了奇偶检验的冗余性盘结构（RAID4、5）的情况．图 14 示出使用本发明构成的 RAID5 结构的盘存储装置的原理图．这是在图 1 的结构之上添加了用于赋予冗余性的盘 25 的结构，控制装置 1、盘装置 2（21、22、23、24）、易失性存储器 3、非易失性存储器 4、时间印记 31、写入缓冲器 41 和缓冲器管理表 42 具有与图 1 所示的实施例相同的功能．

对于图 14 所示的实施例的工作，着眼于与图 1 所示的实施例的差别进行说明．在写入处理中，在主机来的写入数据以只比 1 条少 1 个块的数（K×4－1）存储于写入缓冲器时，控制装置 1 将这些数据写入盘装置 21～25 中．这时，在由作为最后的写入块存储在缓冲器管理表 42

中的各块的逻辑地址和易失性存储器 3 上的时间印记 31 作成逻辑地址标记块之前，与图 1 所示的实施例相同。

然后，根据附加了该逻辑地址标记块的 1 个条的数据进行每一个条单元的异或逻辑（XOR）运算，作成奇偶性的条单元。而且，将该带奇偶的条单元的数据整理后同时写入盘装置 21～25 的空闲区内。此外，时间印记 31 的值在写入结束时刻加 1。这样，将很多零碎的写入归纳成 1 次，而且计算奇偶性时不必读出旧数据和旧的奇偶性块，所以，能够进一步减少存取次数。再有，条的重新装入处理的数据写入也一样，在作成带奇偶的条之后写入盘装置 2。该状态示于图 15。

在奇偶 RAID 的结构中，即使 1 台盘装置发生故障，通过计算发生故障的盘的数据和构成条的其他盘的数据的奇偶异或（XOR），可以再现发生故障的盘的数据，可以继续进行作为盘存储装置的服务。但是，当系统启动时 1 个盘发生故障时，因为还要读出没有存储逻辑地址标记的盘装置的数据并在再生逻辑地址标记之后进行检查，所以，作成变换映象很费时间，大大地增加了系统启动所需要的时间。

因此，如图 16 所示，控制成将构成条的数据块减少 1 个，把相同的逻辑地址标记写入 2 个盘装置。由此，即使 1 个盘装置发生故障，因为在作成变换映象时可以读出另一个盘装置的逻辑地址标记，所以，能够避免大幅度增加系统启动所要的时间。

此外，在使用专用标记区来高速地作成变换映象时，如图 17 所示那样，通过控制专用标记区的逻辑地址标记的分配以便使专用标记区中存储逻辑地址标记的盘装置和存储在条中的盘装置不同，使条内的逻辑地址标记只要 1 个就行了。

再有，在向专用标记区写入逻辑地址标记时，若利用奇偶校验去对付盘故障，过去增加 1 次写入即可，但在这里必需要 2 次写入和 2 次读出，这样，大大地增加了一并写入和条的重新装入时的额外的盘写入时间。因此，该专用区的信息不能用奇偶校验来对付故障。该信息是用于使变换映象高速化，存储在有故障的盘装置的专用标记区中的逻辑地址标记也可以看作是条中逻辑地址标记（随机存取时），所以没有问题。此外，因用随机存取检查的逻辑地址标记只有 1/5，故对于高速八成变换映象很有效果。

本发明可以适用于所有的不改写旧数据区而预先保存更新数据并归

纳起来写入到盘装置内的事先准备好的另外的空闲区中的方法是有效的领域．即，主要可以适用于盘装置和 RAID 结构的存储装置，所述盘装置不仅包括磁盘还包括在顺序写入和随机写入方面性能很不相同的光磁盘等，所述存储装置是具有在更新小块时需要 2 次读出和 2 次写入的奇偶校验的冗余性的 RAID 结构的存储装置．

如上所述，若按照本发明，因原理上无论何时都可以再生变换映象，故不必为了防备电源故障而将变换映象保存在非易失性存储器中．因此，可以构筑非常便宜的盘存储装置．此外，当因硬件故障而使非易失性存储器中的内容丢失时，以往的方法因不能再生变换映象故盘上的数据全部丢失，而现在不同，只是保存在写入缓冲器中的最近的写入数据才丢失，而盘上绝大部分数据仍然完好地保留下来了．因此，大大地提高了对抗故障的能力．进而，因从电源故障恢复过来的处理和通常的系统启动处理完全相同，故不需要系统终结和恢复时的特别处理，从而降低了开发成本．

此外，系统启动时的处理，也因能够通过将逻辑地址标记分散配置在多个盘装置中、设置可以顺序存取逻辑地址标记的专用标记区和对存储区的段分割管理等能实现高速化，故能够将系统启动时的等待时间控制在实际使用时不会发生问题的范围内．特别是，在奇偶性的 RAID 结构中，通过将逻辑地址标记记录在 2 个盘装置中，即使 1 个盘装置发生故障也可以不增加系统启动的时间．

说 明 书 附 图

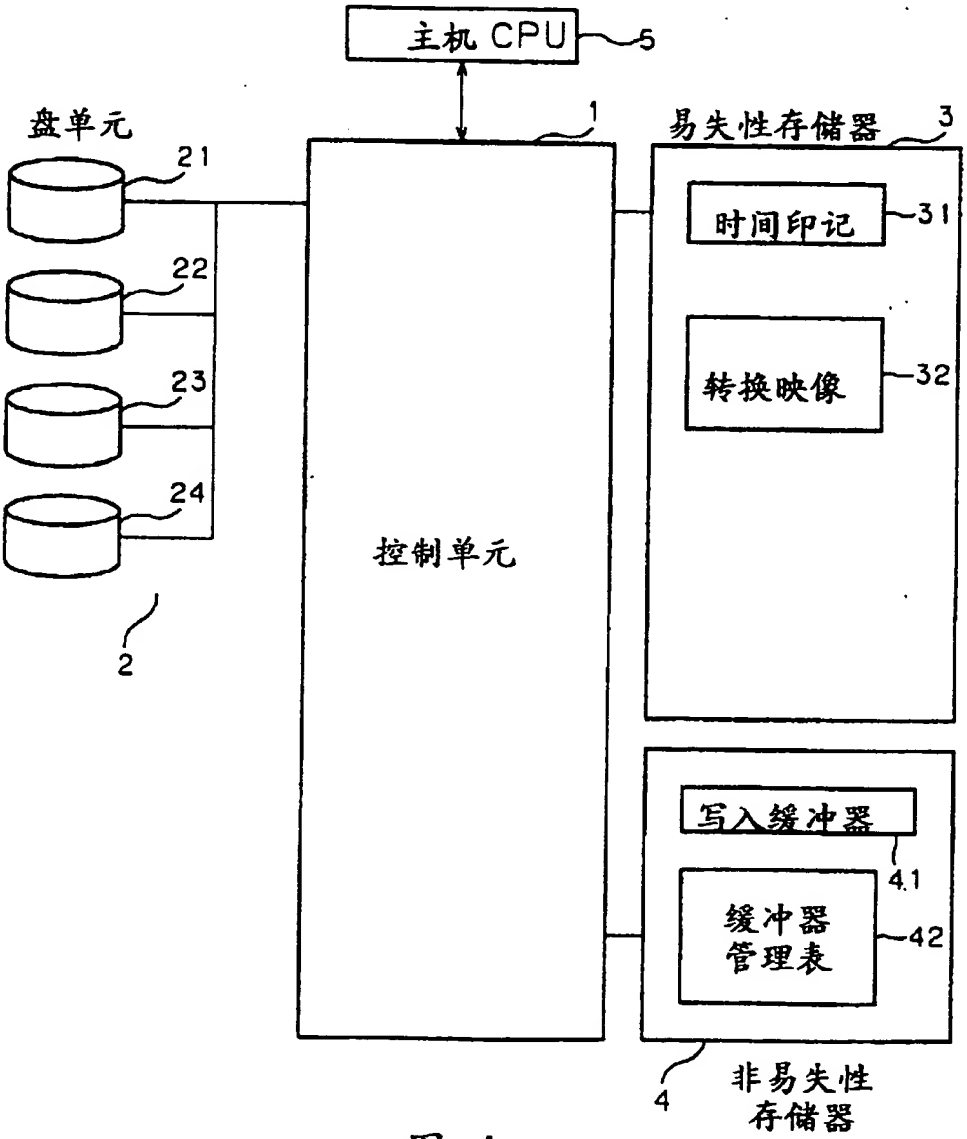

主机 CPU
5
盘单元
1
易失性存储器
3
21
22
23
24
时间印记
31
转换映像
32
控制单元
2
写入缓冲器
41
缓冲器
管理表
42
4
非易失性
存储器

图 1

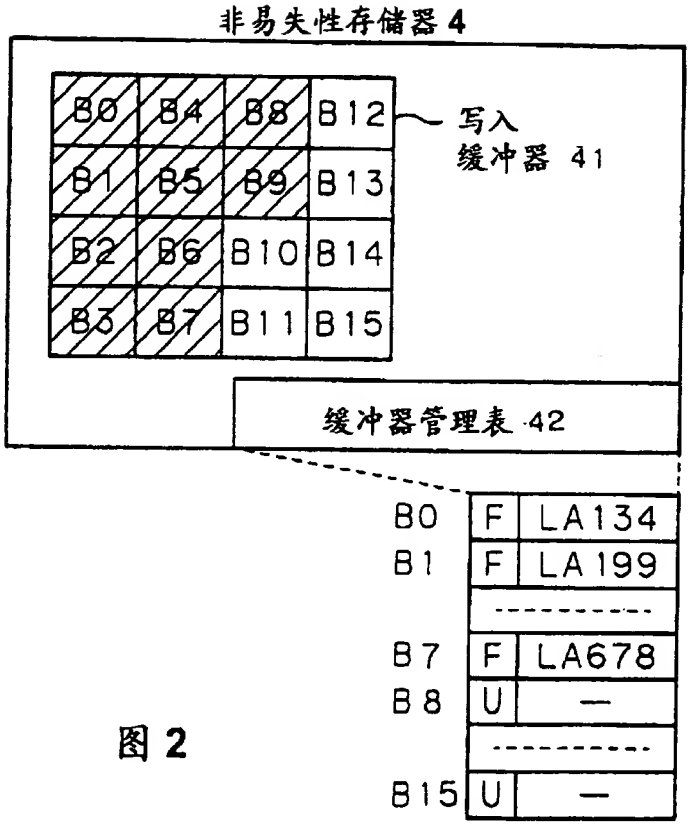

非易失性存储器 4
B0
B4
B8
B12
B1
B5
B9
B13
B2
B6
B10
B14
B3
B7
B11
B15
写入
缓冲器 41
缓冲器管理表 42
B0 F LA134
B1 F LA199
B7 F LA678
B8 U —
B15 U —

图 2

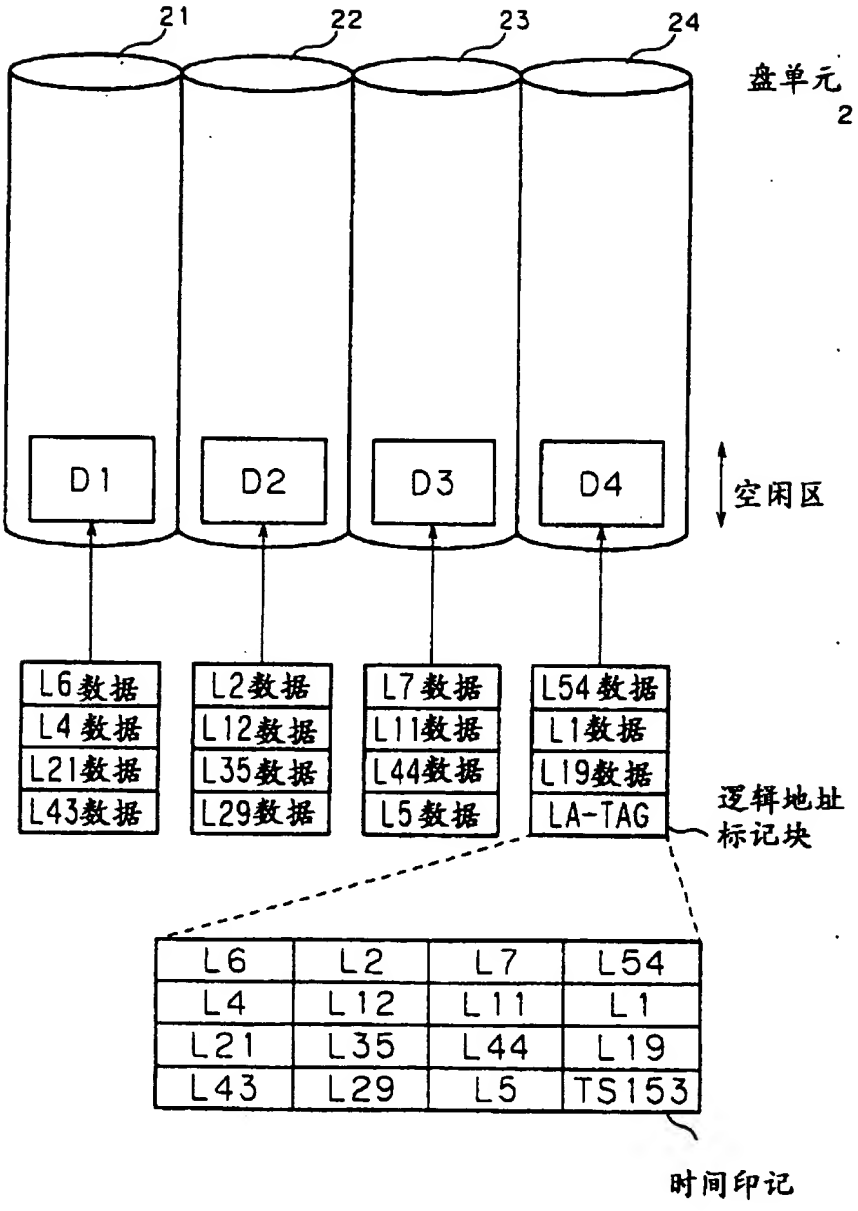

21
22
23
24
盘单元
2
D1
D2
D3
D4
空闲区
L6数据
L4数据
L21数据
L43数据
L2数据
L12数据
L35数据
L29数据
L7数据
L11数据
L44数据
L5数据
L54数据
L1数据
L19数据
LA-TAG
逻辑地址
标记块
L6
L2
L7
L54
L4
L12
L11
L1
L21
L35
L44
L19
L43
L29
L5
TS153
时间印记

图 3

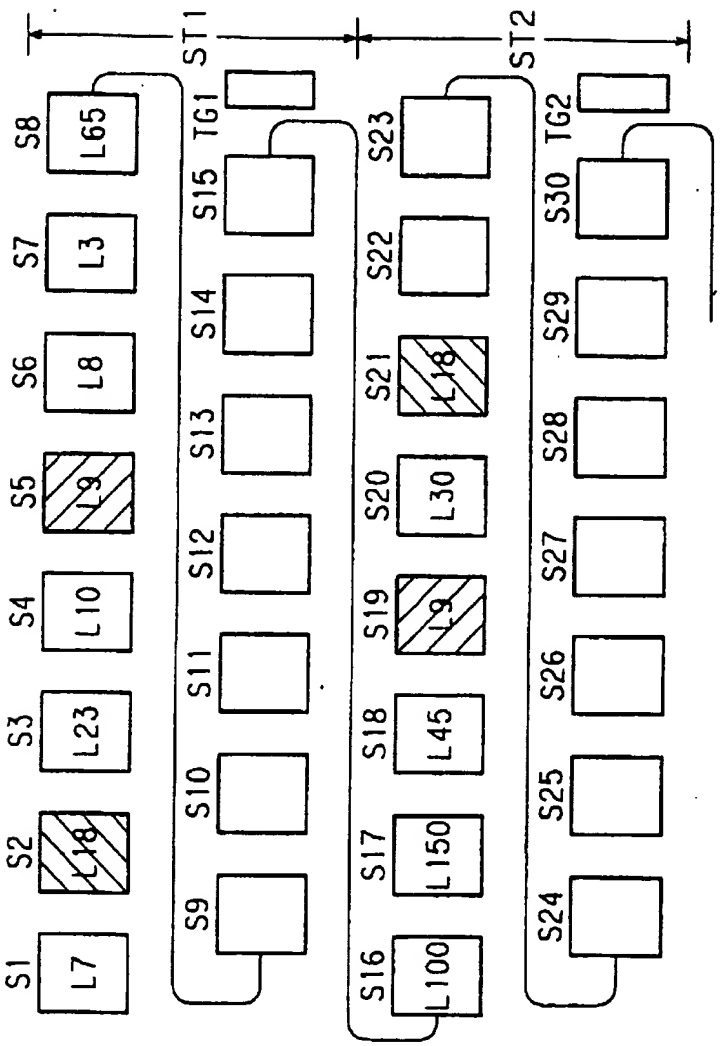

S1
L7
S2
L18
S3
L23
S4
L10
S5
L9
S6
L8
S7
L3
S8
L65
S9
S10
S11
S12
S13
S14
S15
TG1
ST1
S16
L100
S17
L150
S18
L45
S19
L9
S20
L30
S21
L18
S22
S23
ST2
S24
S25
S26
S27
S28
S29
S30
TG2

图 4

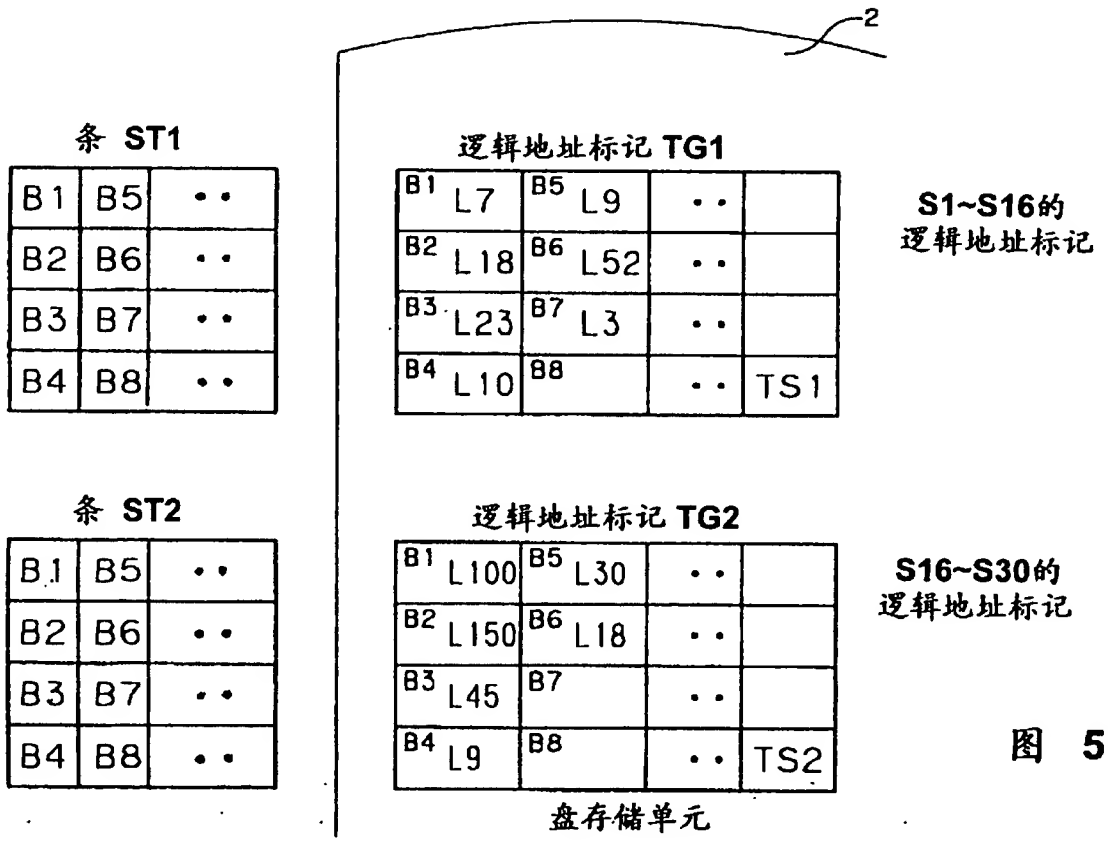

2
条 ST1
B1 B5 ••
B2 B6 ••
B3 B7 ••
B4 B8 ••
逻辑地址标记 TG1
B1 L7 B5 L9 ••
B2 L18 B6 L52 ••
B3 L23 B7 L3 ••
B4 L10 B8 •• TS1
S1~S16的
逻辑地址标记
条 ST2
B1 B5 ••
B2 B6 ••
B3 B7 ••
B4 B8 ••
逻辑地址标记 TG2
B1 L100 B5 L30 ••
B2 L150 B6 L18 ••
B3 L45 B7 ••
B4 L9 B8 •• TS2
S16~S30的
逻辑地址标记
盘存储单元

图 5

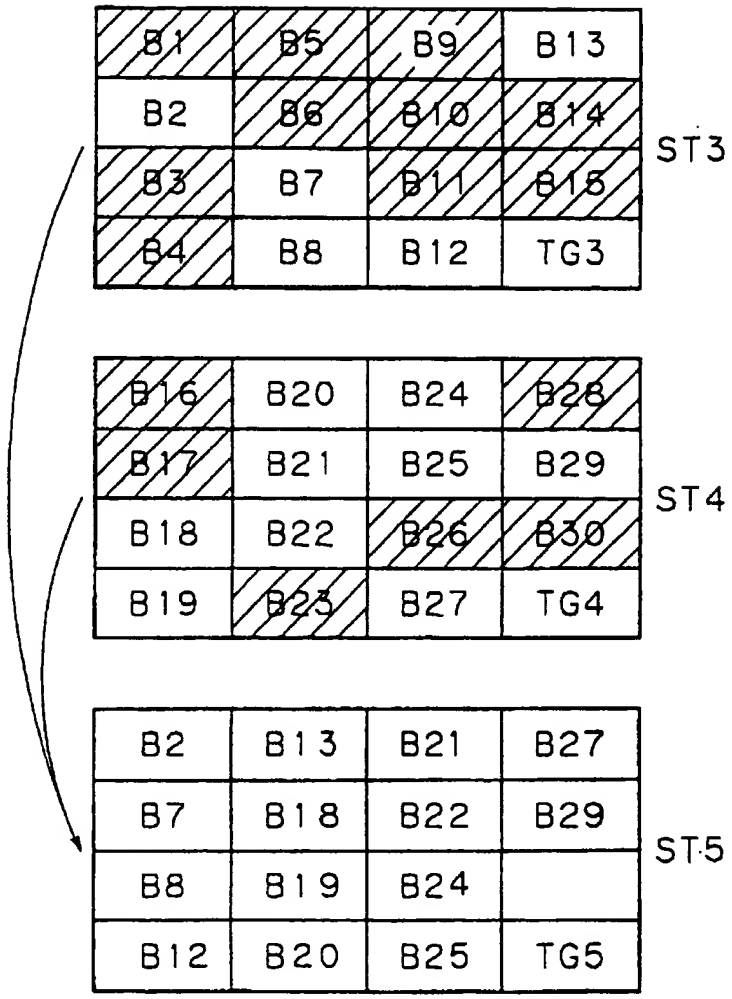

B1
B5
B9
B13
B2
B6
B10
B14
ST3
B3
B7
B11
B15
B4
B8
B12
TG3
B16
B20
B24
B28
B17
B21
B25
B29
ST4
B18
B22
B26
B30
B19
B23
B27
TG4
B2
B13
B21
B27
B7
B18
B22
B29
ST5
B8
B19
B24
B12
B20
B25
TG5

图 6

			L14
L11			
	L13		
	L7	L3	T3

TG3

	L37	L51	
	L41	L46	L55
L25	L22		
L23		L38	T4

TG4

L11	L14	L41	L38
L13	L25	L22	L55
L7	L23	L51	
L3	L37	L46	T8

TG5

图 7

逻辑地址	ST#	BLK#	TS#
L0			
L1			
L2			
⋮			

图8

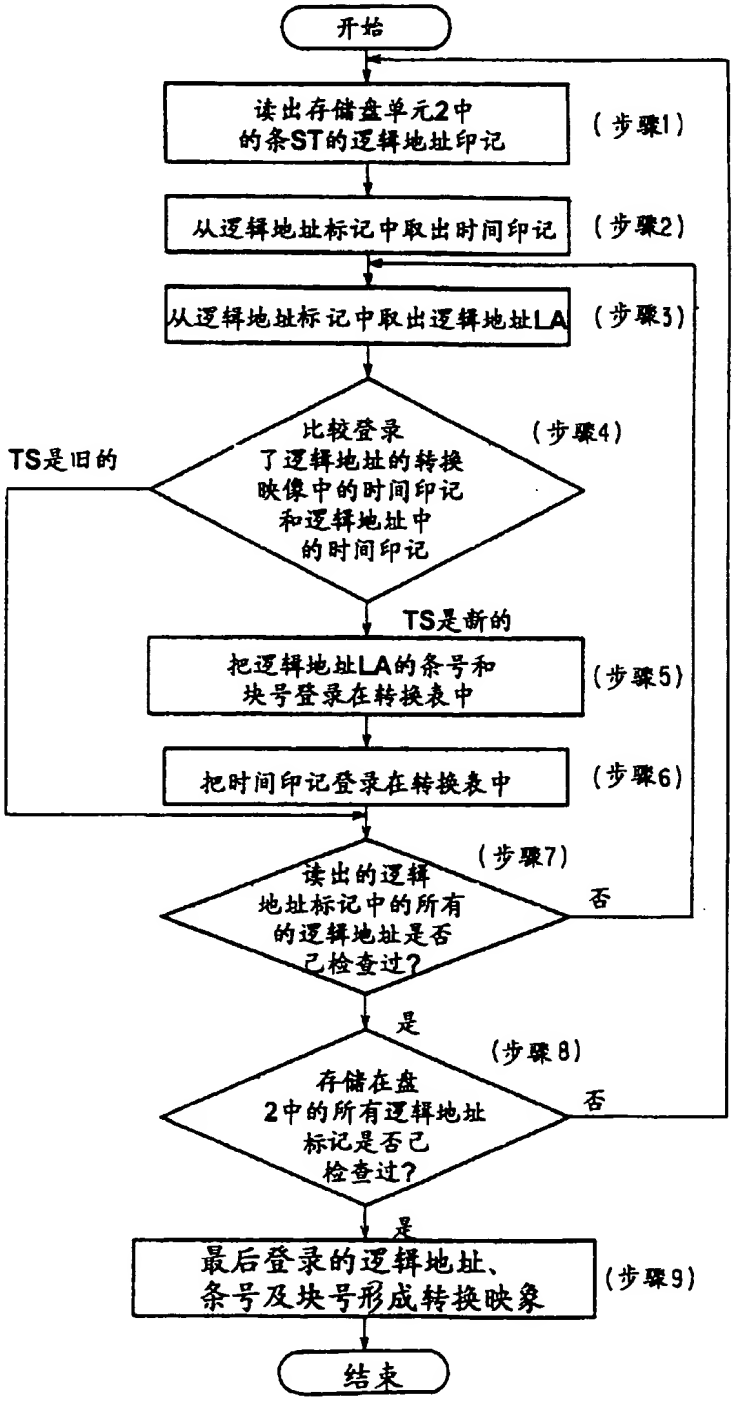

开始
读出存储盘单元2中的条ST的逻辑地址印记
(步骤1)
从逻辑地址标记中取出时间印记
(步骤2)
从逻辑地址标记中取出逻辑地址LA
(步骤3)
比较登录了逻辑地址的转换映像中的时间印记和逻辑地址中的时间印记
(步骤4)
TS是旧的
TS是新的
把逻辑地址LA的条号和块号登录在转换表中
(步骤5)
把时间印记登录在转换表中
(步骤6)
读出的逻辑地址标记中的所有的逻辑地址是否已检查过?
(步骤7)
否
是
存储在盘2中的所有逻辑地址标记是否已检查过?
(步骤8)
否
是
最后登录的逻辑地址、条号及块号形成转换映象
(步骤9)
结束

图 9

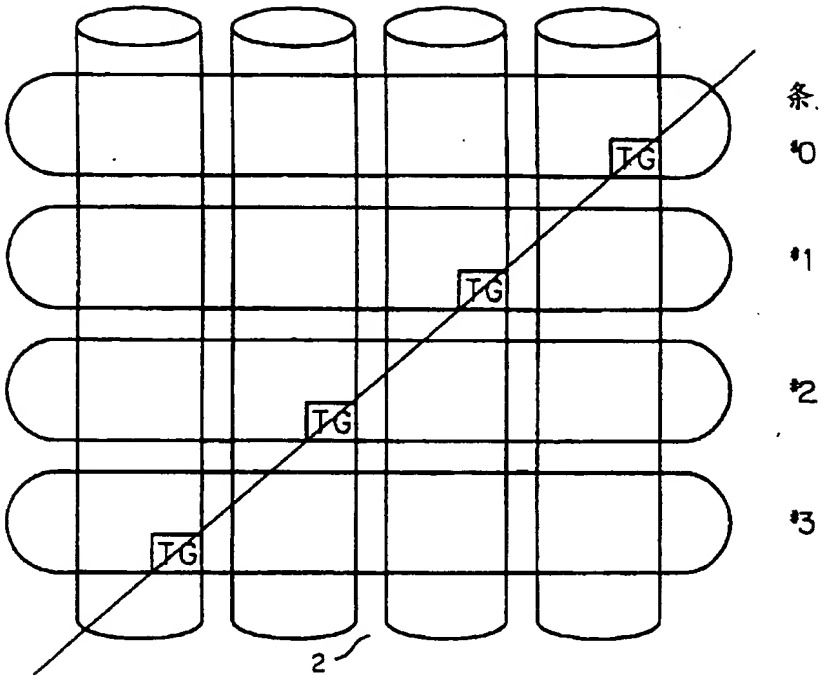


图 10

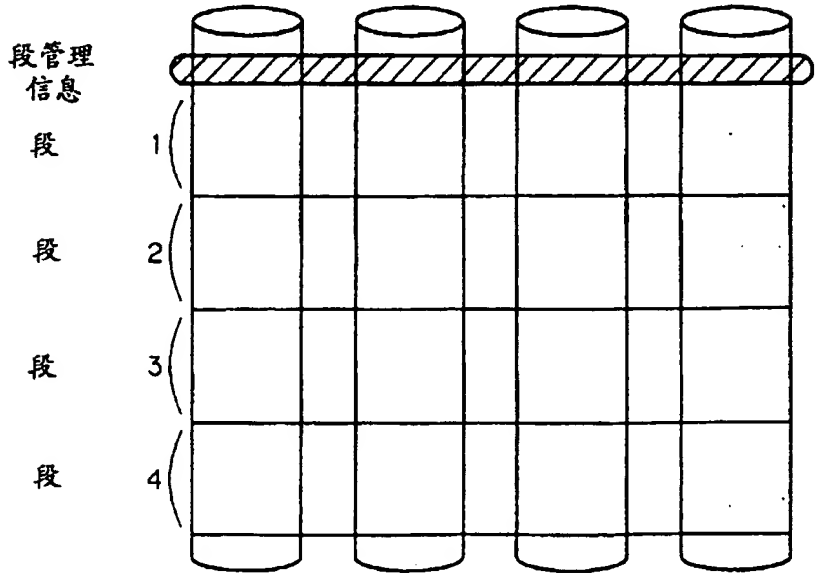


图 11

SG#	转变时的转换映像

段管理信息

图 12

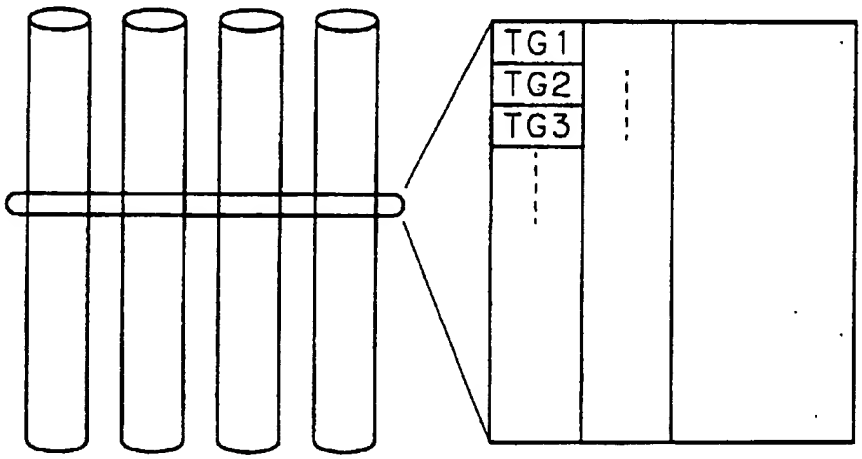


图 13

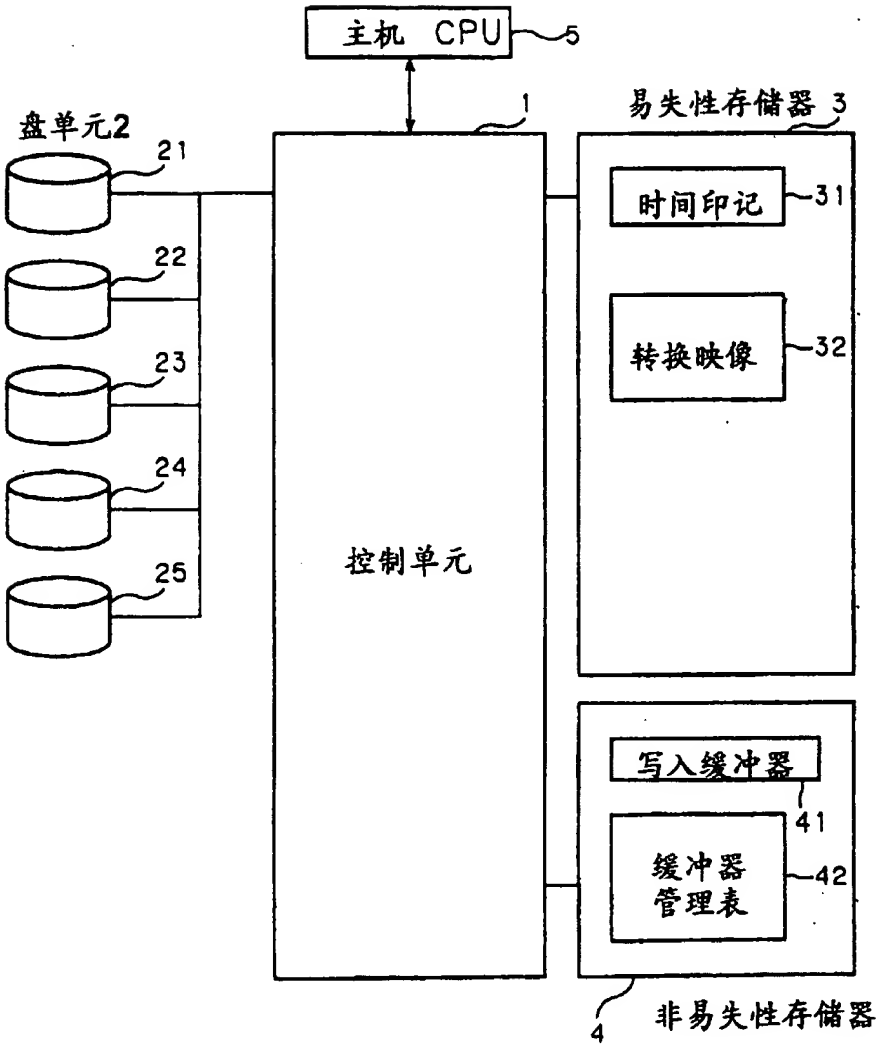

主机 CPU
5
1
易失性存储器 3
盘单元2
21
22
23
24
25
时间印记
31
转换映像
32
控制单元
写入缓冲器
41
缓冲器
管理表
42
4
非易失性存储器

图 14

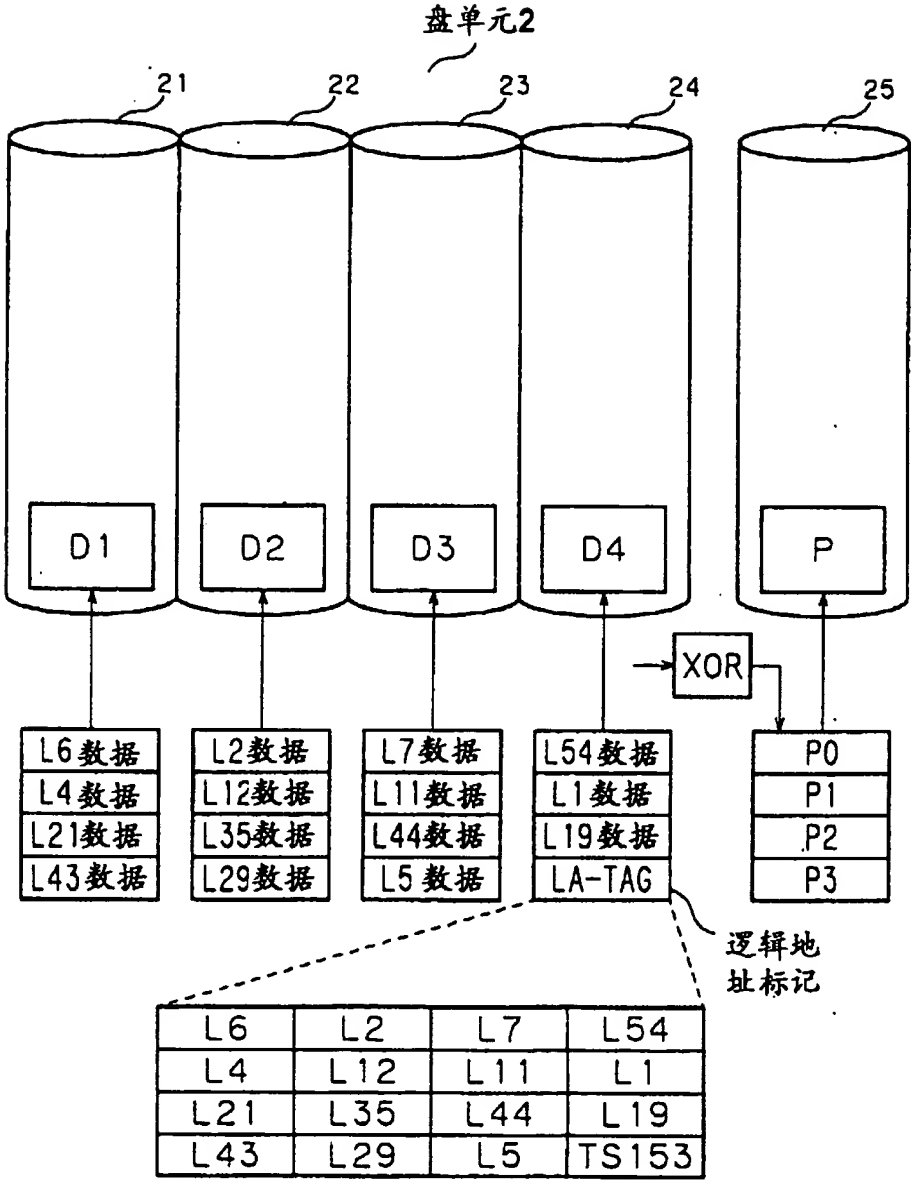

盘单元2
21
22
23
24
25
D1
D2
D3
D4
P
XOR
L6数据
L4数据
L21数据
L43数据
L2数据
L12数据
L35数据
L29数据
L7数据
L11数据
L44数据
L5数据
L54数据
L1数据
L19数据
LA-TAG
P0
P1
P2
P3
逻辑地址标记
L6 L2 L7 L54
L4 L12 L11 L1
L21 L35 L44 L19
L43 L29 L5 TS153

图 15

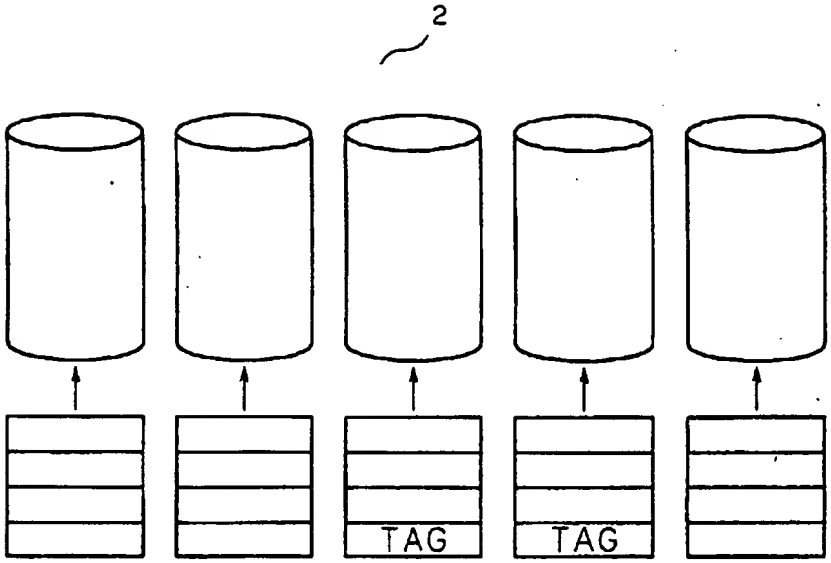

2
TAG
TAG

图 16

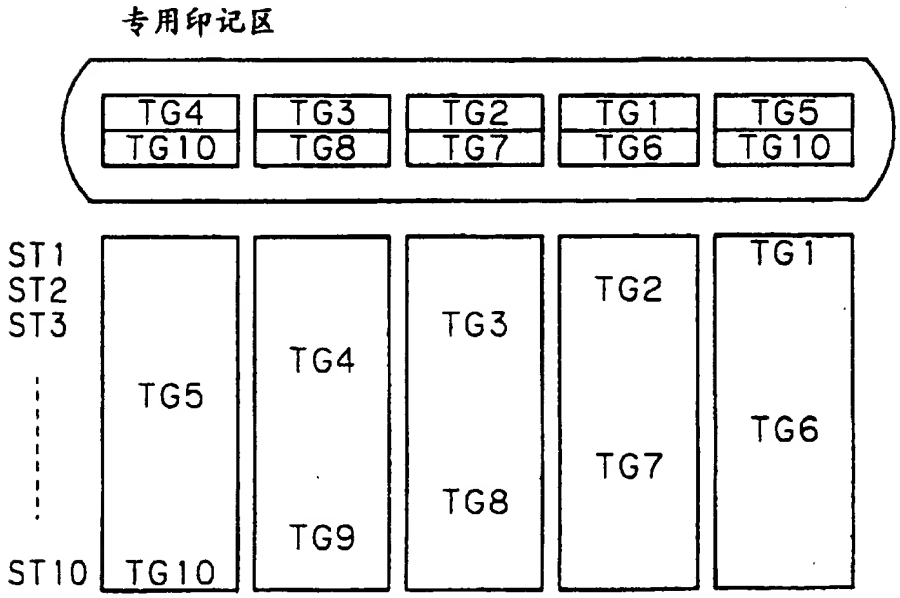

专用印记区
TG4
TG10
TG3
TG8
TG2
TG7
TG1
TG6
TG5
TG10
ST1
ST2
ST3
ST10
TG5
TG10
TG4
TG9
TG3
TG8
TG2
TG7
TG1
TG6

图 17

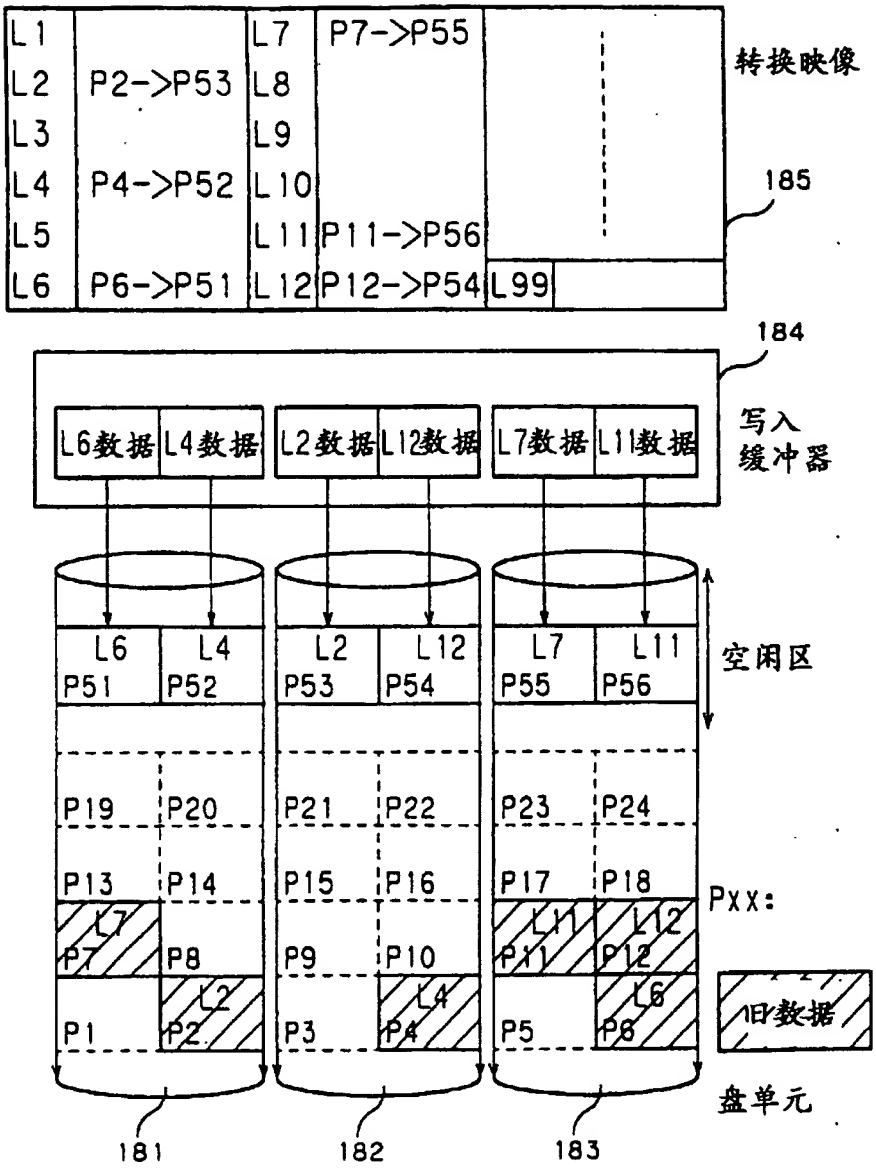

L1
L2 P2->P53
L3
L4 P4->P52
L5
L6 P6->P51
L7 P7->P55
L8
L9
L10
L11 P11->P56
L12 P12->P54
L99
转换映像
185
184
写入
缓冲器
L6数据
L4数据
L2数据
L12数据
L7数据
L11数据
空闲区
L6 L4
P51 P52
L2 L12
P53 P54
L7 L11
P55 P56
P19 P20
P21 P22
P23 P24
P13 P14
P15 P16
P17 P18
Pxx:
L7
P7
P8
P9 P10
L11
P11
L12
P12
L2
P2
P1
L4
P4
P3
P5
L6
P6
旧数据
盘单元
181
182
183

图 18